

# From flat lists to hierarchical domain labels in general language dictionaries

Ana Salgado<sup>1,2</sup>, Toma Tasovac<sup>3</sup>, Rute Costa<sup>1</sup>

<sup>1</sup> NOVA CLUNL, Universidade NOVA de Lisboa, Lisbon, Portugal

<sup>2</sup> Academia das Ciências de Lisboa, Lisbon, Portugal

<sup>3</sup> Belgrade Centre for Digital Humanities, Belgrade, Serbia

ASIALEX 2022

19 June 2022

# Motivation

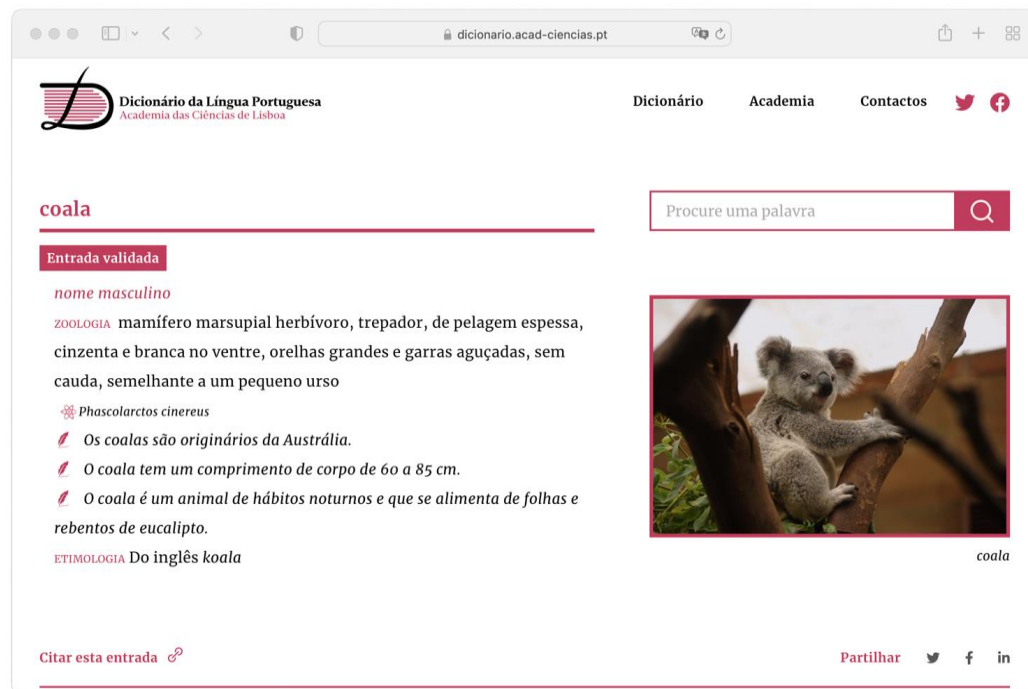
- to improve the lexicographic work carried out on the Lisbon Academy of Sciences (ACL)
- to improve metalanguage and lexicographic description in general
- to guarantee a consistent representation of lexicographic data

# Goals

- an application of TEI Lex-0 specifications to an actual dictionary project – the *Dicionário da Língua Portuguesa* (DLP) from the ACL
- a contribution to the TEI Lex-0 specifications, namely to highlight and discuss the importance of having hierarchical domain labels in TEI also bearing in mind possible future alignments between different lexicographic resources

# Dicionário da Língua Portuguesa (ACL)

**DLP (2022):** *Dicionário da Língua Portuguesa*,  
Academia das Ciências de Lisboa, digital  
(coord. Ana Salgado)



**DLPC (2001):** *Dicionário da Língua Portuguesa Contemporânea*,  
Academia das Ciências de Lisboa, 2 vols.  
(coord. João Malaca Casteleiro)



69 426 entries, 167 556 senses

# Dicionário da Língua Portuguesa (ACL)



## logração

**logração** [logɾasjw̃]. *s. f.* (De *lograr* + sufix. *-ção*). 1. Acção de *lograr*. 2. Engano, logro. = *BURLA*.  
**logrado**, *a* [logɾádu, -te]. *adj.* (Do part. pas. do *v. lograr*). 1. Que foi gozado, desfrutado; que se logrou. 2. Que foi aproveitado. 3. Que foi bem sucedido; que alcançou o seu objectivo. 4. Que foi enganado, burlado.

**logradouro**, **logradouro** [logɾadúru]. *s. m.* (De *lograr* + sufix. *-douro, -douro*). 1. O que é ou pode ser gozado, fruído ou logrado por alguém. 2. Espaço livre, não construído, de uso público. *Os jardins são óptimos logradouros*. 3. Espaço contíguo ou anexo a uma casa e que serve para diversos fins rurais, terreno. 4. Paragem para os gados e que é de uso público.

**lograr** [logɾá]. *v.* (Do lat. *lucrare* 'ter lucro'). 1. Tirar prazer, satisfação. = *APROVEITAR*, *DESFRUTAR*, *FRUIR*, *GOZAR*. 2. Tirar o lucro ou proveito. = *APROVEITAR*. 3. Conseguir fazer alguma coisa; ser bem sucedido. = *ALCANÇAR*, *CONSEGUIR*. ≠ *FALHAR*. *Não logrou passar no exame*. 4. Fazer cair num engano, num logro. = *BURLAR*, *ENGANAR*, *INTROJAR*. *O comerciante pretendia lograr os seus clientes*. 5. Produzir o efeito esperado ou desejado. *A mudança, pelos vistos, logrou*.

**logro** [lógru]. *s. m.* (Do lat. *lucrum* 'lucro'). 0. Acto ou efeito de lograr. 1. Acto de gozar ou fruir. = *FRUIÇÃO*, *GOZO*. 2. *Divers*. Lucro, ganho, proveito. 3. Engano praticado de má-fé contra alguém; acção com carácter fraudulento. = *BURLA*, *DOLO*, *EMBUSTE*, *INTROIÇA*. *O vendedor era um vendedor hábil mas não conseguiu fazê-lo cair no logro de comprar a casa por aquele preço. «É o caso ia esquecendo. E a Acção constituiu-se a Anti-Ação»; a maior mentira, a mais descarada dos logros.* (U. T. RODRIGUES, *Insumissos*, p. 262). 4. Engano com carácter jocoso; partida. 5. Equívoco; confusão. *«Perco-me aos primeiros passos, volto para trás e logo atinjo o logro do meu tremendo suco: agüdeinho não era genze, mas um pobre espanhalês* (R. CORREIA, *Tratado*, p. 70).

**loibé** [lojbe]. *s. m.* *Guin*. 1. Pessoa, geralmente da tribo fula, que tem por profissão construir canoas, a partir de troncos de árvores de grande porte. 2. Artífice que fabrica objectos em madeira.

**loica** *s. f. V. louca.*

**loicaria** *s. f. V. louçaria.*

**loiceira** *s. f. V. louçeira*.

**loiceira** *s. f. V. louçeira*.

**loiceiro** *s. m. V. louçeiro*.

**loiceiro** *s. m. V. louçeiro*.

**loio** *s. m. V. louio*.

**loio** *s. m. V. louio*.

**loio** *s. m. V. louio*.

**loio** *s. m. V. louio*.

**loio** *s. m. V. louio*.

**loio** *s. m. V. louio*.

**loio** *s. m. V. louio*.

**loio** *s. m. V. louio*.

**loio** *s. m. V. louio*.

**loio** *s. m. V. louio*.

**loio** *s. m. V. louio*.

**loio** *s. m. V. louio*.

**loio** *s. m. V. louio*.

**loio** *s. m. V. louio*.

**loio** *s. m. V. louio*.

**loio** *s. m. V. louio*.

**loio** *s. m. V. louio*.

**loio** *s. m. V. louio*.

**loio** *s. m. V. louio*.

**loio** *s. m. V. louio*.

**loio** *s. m. V. louio*.

**loio** *s. m. V. louio*.

**loio** *s. m. V. louio*.

**loio** *s. m. V. louio*.

**loio** *s. m. V. louio*.

**loio** *s. m. V. louio*.

**loio** *s. m. V. louio*.

**loio** *s. m. V. louio*.

**loio** *s. m. V. louio*.

**loio** *s. m. V. louio*.

**loio** *s. m. V. louio*.

**loio** *s. m. V. louio*.

**loio** *s. m. V. louio*.

**loio** *s. m. V. louio*.

**loio** *s. m. V. louio*.

**loio** *s. m. V. louio*.

**loio** *s. m. V. louio*.

## lombada

**loissas, lousas** [lójssɐ], [lójssɐ]. *elem. nom. de loc.* (De *lois*, *lousa*, *coisar* e *loissas*).

**loja** [lɔʒɐ]. *s. f.* (Do fr. *loge*). 1. Piso térreo de um edifício, destinado quer a alojar animais, num meio rural, quer a servir de armazém ou oficina. 2. *Region. (Tr-os-M.)*. Poçilga, parte de um edifício reservada ao alojamento dos porcos. 3. Estabelecimento comercial onde se vendem artigos não consumíveis dentro do próprio estabelecimento. *Na época de Natal, as lojas da Baixa estão muito bem enfeitadas. Ver as maneiras das lojas. loja de adelo*, estabelecimento onde se vendem objectos usados. = *FERRÓVELHO*. *loja de miudezas*, estabelecimento comercial onde se vendem artigos de retrosaria. *loja de modas*, estabelecimento onde se vende especialmente vestuário. 4. Pequena cavidade. 5. *Bot.* Cada uma das subdivisões da antera ou do ovário. 6. *loja maçónica* ou *loja*. 1. Local de reunião de uma associação maçónica. 2. Associação da maçnaria; secção de uma ordem da maçnaria. *loja de pedreiros-livres*, o m. que *loja maçónica*. *desamparar-me/le...*

**loja** [lɔʒɐ]. *s. f.* (De *loja* + sufix. *-ca*). *Deprec.* Estabelecimento de reduzida dimensão onde se exerce uma actividade comercial muito modesta. = *BAIUXA*, *LOCADINA*. *Com o pouco capital que conseguiu reunir abriu, no bairro, uma loja de roupa para criança. «Ao menos ali, naquela loja de espelhos, com flores ricasadas a subão para disfarçar as rachas dos vidros, deixavam-me em paz e respeitavam o meu maxiloso de televisão sintonizado.* (J. GOMES FERREIRA, *Mundo dos Quatro*, p. 160).

**lojeira**, *a* [lɔʒɐjɐ, -ɐ]. *s. f.* (De *loja* + sufix. *-eira*). *Guin*. O m. que *lojista*. *«pedreiros e lojeiros espalhados pelo março entraram de roldão em sequiosa disputa pela baga»* (HERCULANO, *Narrativas e Crónicas*, p. 15).

**lojista** [lɔʒɐjɐ]. *s. m. e f.* (De *loja* + sufix. *-ista*). Proprietário de um estabelecimento comercial; dono de loja. *A associação dos lojistas de Lisboa convocou uma reunião geral de sócios. «Nunca Ernesto ali entrava. Ignorava quem lá moraria. Em troca, conhecia, de vista, quase todos os lojistas em torno do quarteirão; o encadernador, a dona da capelista, o barbeiro, o hingaro da casa de malhas, o cabeleireiro de senhoras, o sapateiro remendado...»* (U. T. RODRIGUES, *Insumissos*, p. 264).

**lolan** [lɔlɐn]. *s. f.* *lingel*. Bambu usado como material de construção na armação do tecto das casas.

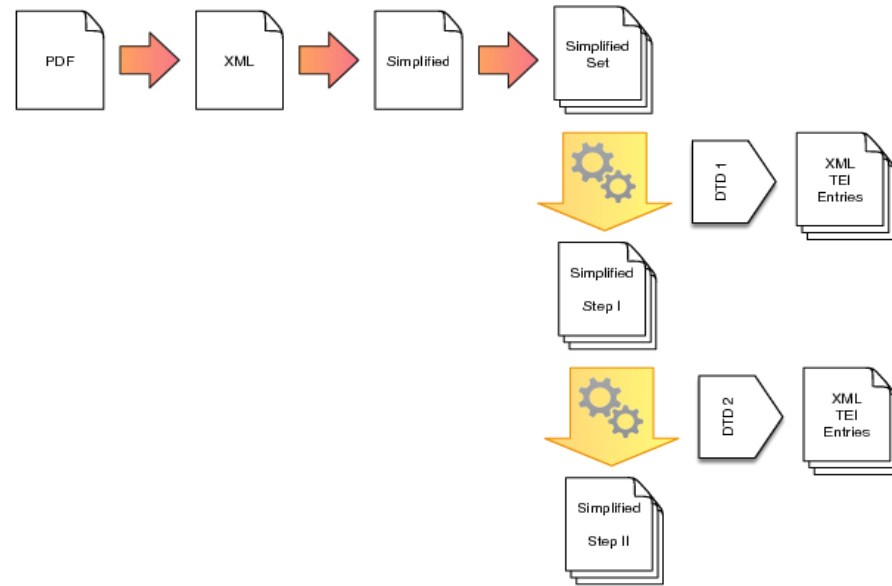
**loliginidas** [lɔlɔʒinidɐs]. *s. m. pl.* (Do lat. *lolligo*, *-nis* 'choco' + sufix. *-ida*). *Zool.* O m. que *loliginídeo*.

**loliginídeo**, *a* [lɔlɔʒinidju, -ɐ]. *adj.* (Do lat. *lolligo*, *-nis* 'choco' + sufix. *-ádeo*). *Zool.* Que pertence ou é relativo aos loliginídeos.

**loliginídeos** [lɔlɔʒinidju]. *s. m. pl.* (Do lat. *lolligo*, *-nis* 'choco' + sufix. *-ádeo*). *Zool.* 1. Família (*Loliginidae*) de moluscos cefalópodes, com duas brânquias e cinco pares de tentáculos, corpo cilíndrico e afunilado na ponta. *A lula pertence à família dos loliginídeos*. 2. *s. m.* Molusco dessa família.

**lomba** [lɔmbɐ]. *s. f.* (De *lombo*). 1. Cumeada ou parte mais alta de uma elevação de terreno; crista arredondada de monte, colina. *«Ultrapassar numa lomba é uma manobra perigosa, proibida pelo Código da Estrada. «Lá no alto, quase na lomba, alcançava-se uma fazenda que fora ourora o brinco da frequência.* (M. BRAGA, *Quatro Irmãs*, p. 150). 2. Pequeno monte de terra ou de areia, natural ou formado por acção do vento; médão. 3. Parte mais elevada de um telhado. 4. *Region. (Tr-os-M.) Pam.* Preguiça, indolência lazeira. 5. *Bras.* Inclinação dos pequenos montes ou das conchilhas de pouca elevação. = *ENCOSTA*, *LOMBADA*.

**lombada** [lɔmbɐdu]. *s. f.* (De *lombo* + sufix. *-ada*). 1. Lomba ou parte mais alta, contínua e prolongada de uma elevação de terreno; grande lomba. 2. Dorso do boi ou de qualquer outro quadrúpede corpulento, em todo o seu



Simões, Almeida & Salgado (2016). Building a Dictionary using XML Technology

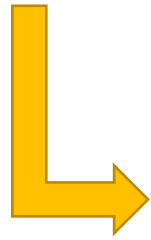
# Dicionário da Língua Portuguesa (ACL)

<b>ACADEMY DICTIONARIES</b>	
<b>Method</b>	semasiological
<b>Normativity</b>	prescriptive in different degrees
<b>Temporal perspective</b>	contemporary
<b>Number of languages</b>	monolingual
<b>Format</b>	structured data formats
<b>Medium</b>	print + digital

*Classification of the Academy Dictionaries under study (Salgado, 2021)*

# Domain labelling

**Usage labelling** = the feature or procedure that provides specific information and is used to mark a particularity of use, **restricting** and **conditioning** the non-regular use of a lexical item



the labels assigned to specialised senses are called **domain labels**



markers that identify the **specialised field of knowledge** in which a lexical unit is mainly used  
Salgado, Costa and Tasovac (2019)

# Flat (non-hierarchical) domain lists

## CLASSIFICAÇÃO DO VOCABULÁRIO QUANTO À REPARTIÇÃO POR CIÊNCIAS, TÉCNICAS E FORMAS DE ACTIVIDADE

**a**

*Acúst.* = Acústica.  
*Aeron.* = Aeronáutica.  
*Agr.* = Agricultura.  
*Álg.* = Álgebra.  
*Alveit.* = Alveitaria.  
*Alven.* = Alvenaria.  
*Anat.* = Anatomia.  
*Antr.* = Antropologia.  
*Apic.* = Apicultura.  
*Arit.* = Aritmética.  
*Arm.* = Armaria.  
*Arqueol.* = Arqueologia.  
*Arquit.* = Arquitectura.  
*Artilh.* = Artilharia.  
*Astr.* = Astronomia.  
*Astronáut.* = Astronáutica.  
*Astrol.* = Astrologia.  
*Autom.* = Automobilismo.

**b**

*Bact.* = Bacteriologia.  
*Balíst.* = Balística.  
*B.-Art.* = Belas-Artes.  
*Biol.* = Biologia (Citologia, Histologia).  
*Bot.* = Botânica.  
*Bromat.* = Bromatologia.

**c**

*Des.* = Desenho.  
*Desp.* = Desporto.  
*Diplom.* = Diplomática.  
*Dir.* = Direito.  
*Dir. Can.* = Direito Canónico.  
*Dir. Civil* = Direito Civil.  
*Dir. Comerc.* = Direito Comercial.  
*Dir. Fiscal* = Direito Fiscal.  
*Dir. Intern.* = Direito Internacional.  
*Dir. Marít.* = Direito Marítimo.

**e**

*Ecl.* = Eclesiástico.  
*Econ.* = Economia.  
*Econ. Pol.* = Economia Política.  
*Electr.* = Electricidade.  
*Electrotéc.* = Electrotécnica.  
*Embr.* = Embriologia.  
*Encad.* = Encadernação.  
*Eng.* = Engenharia.  
*Equit.* = Equitação.  
*Esc.* = Escolar.  
*Escol.* = Escolástica.  
*Escult.* = Escultura.  
*Esg.* = Esgima.  
*Espir.* = Espiritualismo.  
*Estát.* = Estática.  
*Ética.*  
*Etnog.* = Etnografia.

Fragment of the DLPC list

■ Acúst. Acústica	■ Aeron. Aeronáutica	■ Agr. Agricultura	■ Agron. Agronomia
■ Álg. Álgebra	■ Alveit. Alveitaria	■ Alven. Alvenaria	■ Anat. Anatomia
■ Antr. Antropologia	■ Apic. Apicultura	■ Arit. Aritmética	■ Arm. Armaria
■ Arqueol. Arqueologia	■ Arquit. Arquitectura	■ Artilh. Artilharia	■ Astr. Astronomia
■ Astrol. Astrologia	■ Astronáut. Astronáutica	■ Autom. Automobilismo	■ B.-Art. Belas-Artes
■ Bact. Bacteriologia	■ Balíst. Balística	■ Biol. Biologia (citologia, histologia)	■ Bioquim. Bioquímica
■ Bot. Botânica	■ Bromat. Bromatologia	■ Carn. Carniçaria	■ Carp. Carpintaria
■ Cartog. Cartografia	■ Cerâm. Cerâmica	■ Chapel. Chapelaria	■ Cibern. Cibernética
■ Cineg. Cinegética	■ Cinem. Cinema, cinematografia	■ Cirurg. Cirurgia	■ Comérc. Comércio
■ Constr. Construção	■ Contab. Contabilidade	■ Corr. Correios	■ Cosmol. Cosmologia
■ Cristalol. Cristalografia	■ Cronol. Cronologia	■ Cul. Culinária	■ Curt. Curtumes
■ Cutel. Cutelaria	■ Des. Desenho	■ Desp. Desporto	■ Diplom. Diplomática
■ Dir. Direito	■ Dir. Can. Direito Canónico	■ Dir. Civil Direito Civil	■ Dir. Comerc. Direito Comercial
■ Dir. Fiscal Direito Fiscal	■ Dir. Intern. Direito Internacional	■ Dir. Marít. Direito Marítimo	■ Ecles. Eclesiástico
■ Ecol. Ecologia	■ Econ. Economia	■ Econ. Pol. Economia Política	■ Electr. Electricidade
■ Electrotéc. Electrotécnica	■ Embr. Embriologia	■ Encad. Encadernação	■ Eng. Engenharia
■ Equit. Equitação	■ Esc. Escolar	■ Escol. Escolástica	■ Escult. Escultura
■ Esg. Esgima	■ Espir. Espiritualismo	■ Estat. Estática	■ Ética Ética
■ Etnog. Etnografia	■ Étnol. Etnologia	■ Farm. Farmácia	■ Filol. Filologia
■ Filos. Filosofia	■ Fin. Finanças	■ Fis. Física	■ Fis. Atóm. Física Atómica
■ Fisiol. Fisiologia	■ Fonét. Fonética	■ Fort. Fortificação	■ Fot. Fotografia
■ Fut. Futebol	■ Futur. Futurologia	■ Geneal. Genealogia	■ Genét. Genética
■ Geod. Geodesia	■ Geog. Geografia	■ Geol. Geologia	■ Geom. Geometria
■ Gin. Ginástica	■ Gram. Gramática	■ Heráld. Heráldica	■ Hidrául. Hidráulica
■ Hidrog. Hidrografia	■ Hig. Higiene	■ Hist. História	■ Hist. Nat. História Natural
■ Hist. Polit. História Política	■ Histol. Histologia	■ Hort. Horticultura	■ Ind. Indústria
■ Inform. Informática	■ Jard. Jardinagem	■ Jur. Jurídico, jurisprudência	■ Ling. Linguística
■ Literat. Literatura	■ Liturg. Liturgia	■ Lóg. Lógica	■ Magn. Magnetismo
■ Mar. Marítima	■ Marc. Marcenaria	■ Mat. Matemática	■ Mecân. Mecânica
■ Med. Medicina	■ Med. Leg. Medicina Legal	■ Metal. Metalurgia	■ Meteor. Meteorologia
■ Métr. Métrica	■ Metrol. Metrologia	■ Mil. Militar	■ Miner. Mineralogia
■ Mitol. Mitologia	■ Mús. Música	■ Náut. Náutica	■ Neur. Neurologia
■ Numism. Numismática	■ Ocul. Oculatismo	■ Olar. Olaria	■ Ópt. Óptica
■ Ouriv. Ourivesaria	■ Paleog. Paleografia	■ Paleont. Paleontologia	■ Parasit. Parasitologia
■ Pat. Patologia	■ Pec. Pecuária	■ Pedag. Pedagogia	■ Pesc. Pesca, pescaria
■ Pint. Pintura	■ Pirotec. Pirotecnia	■ Pisc. Piscicultura	■ Políc. Polícia
■ Polit. Política	■ Pré-hist. Pré-história	■ Psican. Psicanálise	■ Psicofisiol. Psicofisiologia
■ Psicol. Psicologia	■ Psiq. Psiquiatria	■ Quím. Química	■ Quím. Biol. Química Biológica
■ Quím. Org. Química Orgânica	■ Radiol. Radiologia	■ Rel. Religião	■ Retór. Retórica
■ Sal. Salinas	■ Sapat. Sapataria	■ Serr. Serralharia	■ Serr. Serralharia
■ Sider. Siderurgia, Siderotecnia	■ Silv. Silvicultura	■ Sociol. Sociologia	■ Tanoar. Tanoaria
■ Taurom. Tauromaquia	■ Teat. Teatro	■ Tecnol. Tecnologia	■ Telecom. Telecomunicações
■ Telef. S. Fios Telefonia sem fios	■ Teleg. Telegrafia	■ Teol. Teologia	■ Teos. Teosofia
■ Terat. Teratologia	■ Tip. Tipografia	■ Topog. Topografia	■ Transp. Transportes
■ Trigon. Trigonometria	■ Veloc. Velocipedia	■ Ven. Venatório	■ Veter. Veterinária
■ Vin. Vinificação	■ Vit. Viticultura	■ Zool. Zoologia	■ Zootec. Zootecnia

Domain labels in the DLPC (184)



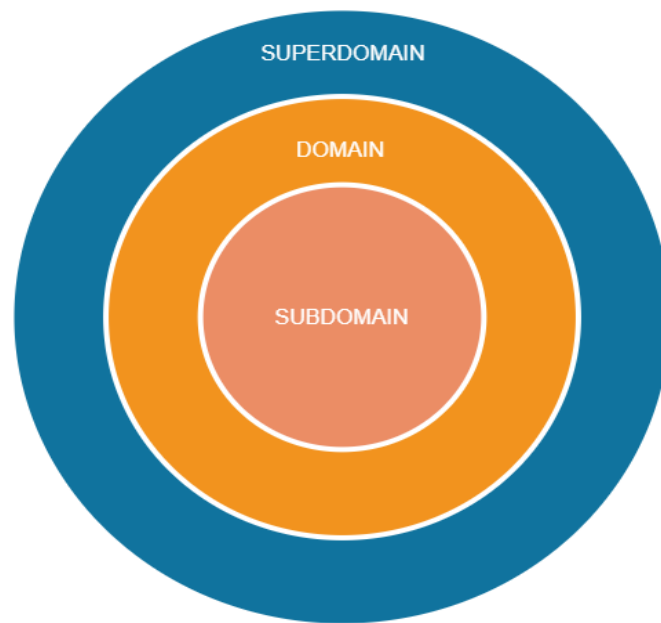


# Flat (non-hierarchical) domain lists

DLPC	DLE	DAF	METALABEL	Dewey Decimal Classification (DDC)	Universal Decimal Classification (UDC)	EuroVoc	UNESCO Thesaurus
cristalografia			<b>crystallography</b>	540 Chemistry & allied sciences/548 Crystallography	54 Chemistry. Crystallography. Mineralogy/548/549 Mineralogical sciences. Crystallography. Mineralogy	<i>crystallography returned 0 results</i>	36 Science/2.20 Physical sciences/Cristallography
geologia	geología	géologie	<b>geology</b>	550 Earth sciences/551 Geology, hydrology & meteorology	55 Earth sciences. Geological sciences / 551 General geology. Meteorology. Climatology. Historical geology. Stratigraphy. Paleogeography	36 SCIENCE/3606 natural and applied sciences/NT1 geology	2 Science/2.35 Earth sciences/Geology
mineralogia		minéralogie	<b>mineralogy</b>	540 Chemistry & allied sciences/549 Mineralogy	54 Chemistry. Crystallography. Mineralogy/548/549 Mineralogical sciences. Crystallography. Mineralogy	36 SCIENCE/3606 natural and applied sciences/NT1 geology/NT2 mineralogy	2 Science/2.35 Earth sciences/Mineralogy
paleontologia		paléontologie	<b>palaeontology</b>	560 Paleontology; paleozoology	56 Paleontology	X	2 Science/2.35 Earth sciences/Paleontology

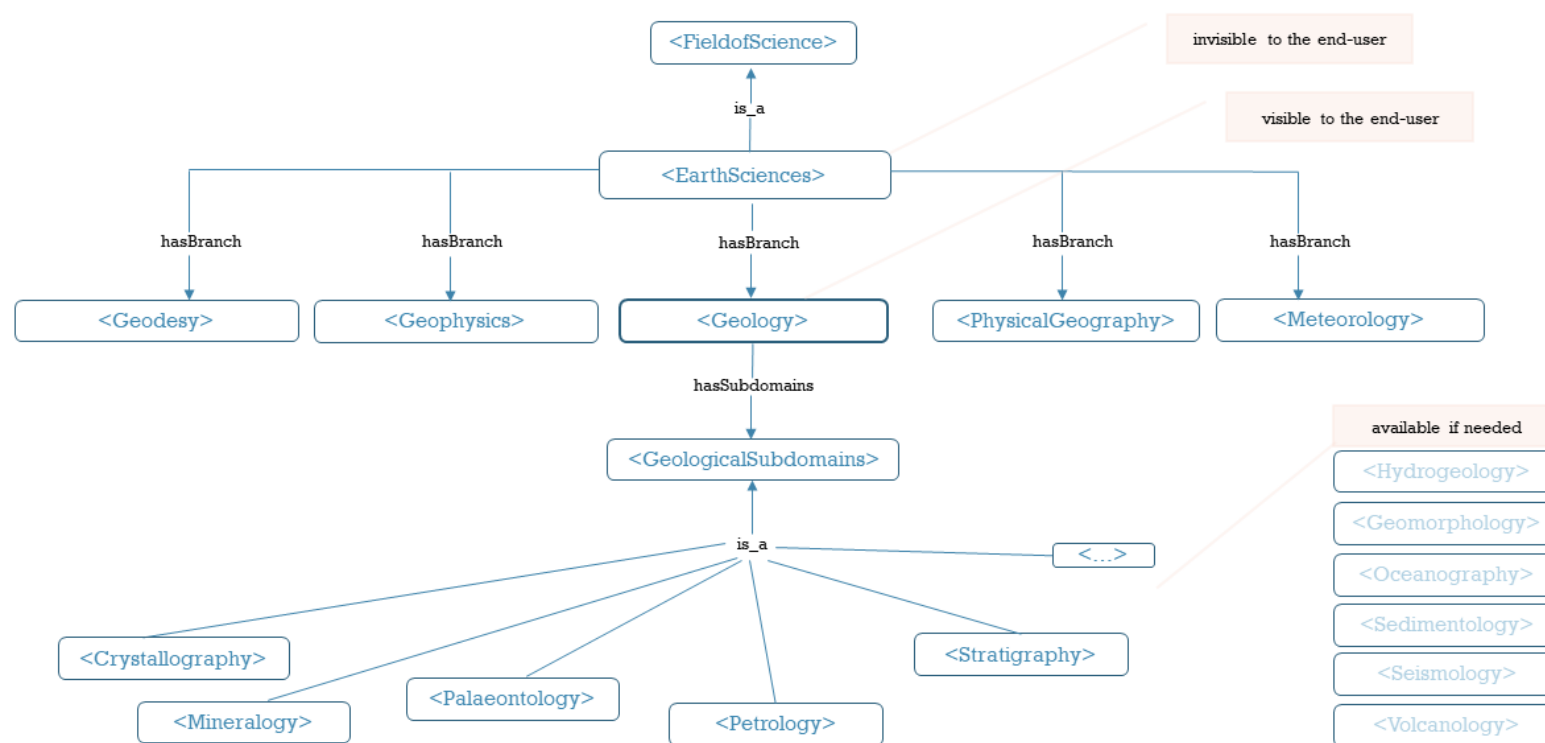
*Comparison of academy dictionaries domain labels and classification systems*  
Salgado, Costa & Tasovac (2021)

# Hierarchical domain labels



- to organise an increasing amount of terminological data
- to provide the lexicographers greater control over specialised content

# Hierarchical domain labels



*Domain labels within the EARTH SCIENCES superdomain showing GEOLOGY as domain and identifying its subdomains*

# Text Encoding Initiative (TEI)

- TEI is the dominant format for several lexicographic projects.
- The TEI Guidelines have a module explicitly focused on the encoding of dictionaries.
- However, this module is criticised regarding its extreme flexibility, i.e., the existence of multiple possibilities to encode similar structures that affect the interoperability of the encoded formats.
- This flexibility is also the characteristic that justifies its wide adoption.
- To reduce this freedom and define a specific format for dictionaries, the **TEI Lex-0** initiative, a stricter version of the TEI schema, has been promoted to reduce the encoding options and to improve interoperability.

*TEI Lex-0 aims at establishing a target format to facilitate the interoperability of heterogeneously encoded lexical resources.*

ROMARY & TASOVAC (2018)

# TEI Lex-0

Restricting the scope of usage information

TEI P5	TEI Lex-0	TEI Lex-0 (reformulation)	Examples
<b>geo</b>	geographic	geo	España, regionalismo
<b>time</b>	temporal	time	anticuado, néologisme
<b>dom</b>	domain	domain	deportes, medicina
<b>reg</b>	socioCultural	socioCultural	calão, jerga
<b>reg</b>		register	familiar, formal
<b>style</b>	meaningType	meaningType	figurative
<b>plev</b>	frequency	frequency	desusado, poco usado
-	attitude	attitude	irónico, peyorativo
-	normativity	normativity	non-standard
<b>style</b>	-	textType	littéraire, poético

**<usg type="domain" />**

# Flat representation of domain labels

```
<usg subtype="superdomain" expand="Ciências da Terra">C. Terra</usg>  
<usg type="domain" expand="Geologia">Geol.</usg>  
<usg subtype="subdomain" expand="Mineralogia">Min.</usg>
```



# Hierarchical domain labels

## *Multilingual hierarchy of EARTH SCIENCES superdomain*

```
<encodingDesc>
  <classDecl>
    <taxonomy xml:id="domain">
      <category xml:id="domain.earth_sciences">
        <catDesc xml:lang="en">Earth Sciences</catDesc>
        <catDesc xml:lang="pt">Ciências da Terra</catDesc>
        <catDesc xml:lang="es">Ciencias de la Tierra</catDesc>
        <catDesc xml:lang="fr">sciences de la Terre </catDesc>
        <category xml:id="domain.earth_sciences.geology">
          <catDesc xml:lang="en">Geology</catDesc>
          <catDesc xml:lang="pt">Geologia</catDesc>
          <catDesc xml:lang="es">Geología</catDesc>
          <catDesc xml:lang="fr">Geologie</catDesc>
          <category xml:id="domain.earth_sciences.geology.mineralogy">
            <catDesc xml:lang="en">Mineralogy</catDesc>
            <catDesc xml:lang="pt">Mineralogia</catDesc>
            <catDesc xml:lang="es">Mineralogía</catDesc>
            <catDesc xml:lang="fr">Mineralogie</catDesc>
          </category>
        </category>
      </category>
    </taxonomy>
  </classDecl>
</encodingDesc>
```



```
<usg type="domain" corresp="#domain.earth_sciences.geology"/>
```



# Summary

- We found advantages in using hierarchical domain labels (superdomain, domain, subdomain). For this, we first need to include a <taxonomy> in the <teiHeader> and then use correspondences designations.
- Use the element <usg> to annotate data about domain labels.
- Assign the "domain" value to the attribute @type.
- If the data uses abbreviated forms, we recommend providing the full form using the @expand attribute if using flat <usg> labels. Encoding the abbreviation and its respective full form at the same time is very useful. Later, we can decide how this information will be viewed when publishing the digital or printed data. If using a taxonomy in the <teiHeader>, full forms should be provided as values in <catDesc> elements.
- The domain label can be associated at various points in the entry hierarchy. Its position must be analysed and evaluated, on a case-by-case basis, by the lexicographer.

# Concluding remarks

- Domain organisation and a domain hierarchy are crucial to improve the labelling system in dictionaries.
- A well-organised domain hierarchy will help lexicographers and end-users better understand the relations between concepts.
- By proposing hierarchical domain labels, we organise knowledge and establish higher and lower categories.
- The fact that we define a domain hierarchy does not mean that all proposed labels will be visible in the final product.
- The solutions presented for the DLP can be replicated.
- The need to apply standardised models within the lexicographic universe reveals that these cannot be closed models.
- The currently recommended TEI Lex-0 practice of representing domain labels as flat values is not robust enough to deal with hierarchical domain structures.
- By observing and comparing various lexicographic resources, we believe we are taking an essential step towards a possible homogenisation of the representation of lexicographic data.

# Directions for future research

- To continue to study Academy dictionaries in more detail.
- To test the proposed methodology on terms from other fields.
- We are interested in exploring the results in the field of ontology.
- To solve the problematic issues found (domains with multiple labels, unlabelled equivalent headwords, outdated definitions, etc.).
- To try to establish an agreement between academies/institutions to systematise and optimise a new type of lexicography that can better represent the entire European lexicographic heritage.
- We will continue to invest in approaches that combine terminological and lexicographic assumptions, and even other disciplines, placing best practice standards at the core of our research.

# Thanks for your attention!

[anasalgado@campus.fcsh.unl.pt](mailto:anasalgado@campus.fcsh.unl.pt)

[ttasovac@humanistika.org](mailto:ttasovac@humanistika.org)

[rute.costa@fcsh.unl.pt](mailto:rute.costa@fcsh.unl.pt)